

0590
0429

13



OIPÉ

RAW SEQUENCE LISTING

DATE: 05/07/2002

PATENT APPLICATION: US/09/874,162A

TIME: 14:55:00

Input Set : A:\05311-024001.txt

Output Set: N:\CRF3\05072002\I874162A.raw

ENTERED

```

4 <110> APPLICANT: Koontz, Jason
5     Sklar, Jeffrey
7 <120> TITLE OF INVENTION: FUSION OF JAZF1 AND JJAZ1 GENES IN
8     ENDOMETRIAL STROMAL TUMORS
10 <130> FILE REFERENCE: 05311-024001
12 <140> CURRENT APPLICATION NUMBER: US 09/874,162A
13 <141> CURRENT FILING DATE: 2001-06-04
15 <150> PRIOR APPLICATION NUMBER: US 60/209,093
16 <151> PRIOR FILING DATE: 2000-06-02
18 <160> NUMBER OF SEQ ID NOS: 23
20 <170> SOFTWARE: FastSEQ for Windows Version 4.0
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 3010
24 <212> TYPE: DNA
25 <213> ORGANISM: Homo sapiens
27 <220> FEATURE:
28 <221> NAME/KEY: CDS
29 <222> LOCATION: (52)...(780)
31 <400> SEQUENCE: 1
32 cccgcccggc gctcgcagag ccgacaccag gggggctctc gatgtagcac c atg aca      57
33                                     Met Thr
34                                     1
36 ggc atc gcc gcc gcc tcc ttc ttc tcc aat acc tgc cga ttc ggg ggc      105
37 Gly Ile Ala Ala Ala Ser Phe Phe Ser Asn Thr Cys Arg Phe Gly Gly
38     5                10                15
40 tgc gga ctc cac ttc ccc acc ctg gcc gac ctc atc gag cac atc gag      153
41 Cys Gly Leu His Phe Pro Thr Leu Ala Asp Leu Ile Glu His Ile Glu
42     20                25                30
44 gac aac cac atc gat aca gat cca cgg gtt tta gaa aaa caa gaa tta      201
45 Asp Asn His Ile Asp Thr Asp Pro Arg Val Leu Glu Lys Gln Glu Leu
46 35                40                45                50
48 cag cag cca acc tat gtt gcc ctg agt tac ata aat aga ttc atg aca      249
49 Gln Gln Pro Thr Tyr Val Ala Leu Ser Tyr Ile Asn Arg Phe Met Thr
50     55                60                65
52 gat gct gcc cgc cga gag cag gag tcc cta aag aag aag att cag ccg      297
53 Asp Ala Ala Arg Arg Glu Gln Glu Ser Leu Lys Lys Lys Ile Gln Pro
54     70                75                80
56 aag ctc tcg ctg act ctg tcc agc tca gtg tct cga ggg aat gtg tcc      345
57 Lys Leu Ser Leu Thr Leu Ser Ser Ser Val Ser Arg Gly Asn Val Ser
58     85                90                95
60 act ccc cca cgc cac agc agt gga agc ctt act ccc ccc gtg acc cca      393
61 Thr Pro Pro Arg His Ser Ser Gly Ser Leu Thr Pro Pro Val Thr Pro
62     100                105                110

```

RAW SEQUENCE LISTING

DATE: 05/07/2002

PATENT APPLICATION: US/09/874,162A

TIME: 14:55:00

Input Set : A:\05311-024001.txt

Output Set: N:\CRF3\05072002\I874162A.raw

```

64 ccc atc acc ccc tcc tct tca ttc cgc agc agc act ccg aca ggc agc      441
65 Pro Ile Thr Pro Ser Ser Ser Phe Arg Ser Ser Thr Pro Thr Gly Ser
66 115                      120                      125                      130
68 gag tat gac gag gag gag gtg gac tat gag gag tcg gac agc gat gag      489
69 Glu Tyr Asp Glu Glu Glu Val Asp Tyr Glu Glu Ser Asp Ser Asp Glu
70                      135                      140                      145
72 tcc tgg acc aca gag agt gcc atc agc tcc gaa gcc atc ctc agc tcc      537
73 Ser Trp Thr Thr Glu Ser Ala Ile Ser Ser Glu Ala Ile Leu Ser Ser
74                      150                      155                      160
76 atg tgc atg aat gga ggg gaa gag aag cct ttt gcc tgc cca gtt cct      585
77 Met Cys Met Asn Gly Gly Glu Glu Lys Pro Phe Ala Cys Pro Val Pro
78                      165                      170                      175
80 gga tgt aaa aag aga tac aag aat gtg aat ggc ata aag tat cac gct      633
81 Gly Cys Lys Lys Arg Tyr Lys Asn Val Asn Gly Ile Lys Tyr His Ala
82                      180                      185                      190
84 aag aat ggt cac aga aca cag att cgt gtc cgc aaa cca ttc aag tgt      681
85 Lys Asn Gly His Arg Thr Gln Ile Arg Val Arg Lys Pro Phe Lys Cys
86 195                      200                      205                      210
88 cgc tgt ggg aag agt tac aag aca gct cag ggc ctg cgg cac cac aca      729
89 Arg Cys Gly Lys Ser Tyr Lys Thr Ala Gln Gly Leu Arg His His Thr
90                      215                      220                      225
92 atc aat ttc cat ccc ccg gtg tcg gct gag att atc agg aag atg cag      777
93 Ile Asn Phe His Pro Pro Val Ser Ala Glu Ile Ile Arg Lys Met Gln
94                      230                      235                      240
96 caa taacatgctg gtcataactg tgccaagaaa tcctcaccag cagttgctga      830
97 Gln
100 ttttgaaaac agccaccttt tttcagggga agcattcagc aaccctttta agaaaaagaa      890
101 ttaaattgcat gcttttaaatt ttttctgtaa ttttggaatg atgtatcttt gtagagttaa      950
102 tgatttttgta catttgacaca tgtaatcatc ataccctatt tcattacttt gatataaggt      1010
103 gctaaacaaa aaaagctcta ggttcttcag cacatttccc ccaaaacaaa ataaaattga      1070
104 gggcatgttg catattgttg aattgtattg cgggtgtatc aacctggggg gaggaggggc      1130
105 tggcactgag attttttttt caagattgta atgtgattga agttttcaac acatcaactc      1190
106 acatatgttc aaaacaaaaa taataccttc attatcaaac tggttaccat gccttacata      1250
107 atggagttag tatttgtgag tagaaagact ttaggtaatg gaaatataaa taagaaagaa      1310
108 tgtttaacat aatatgctaa aaatattttc atattttaa atacatacgt aaggtgtgct      1370
109 ttctgtgttt tatattatct tgcaaatcct tttgcccttt aaaaagctga aaatcttgcc      1430
110 atctgactta ctagtcattt tagtggtata aatggcattt tgtacaaaat agtctattca      1490
111 gttcgttcat tcatttaaca cacattgatt gagtgcctgc tgggtacaag ggattcaatt      1550
112 tatgcctatt gatattctgc gaccaagata cccttttagt gaaatacttt tttccctgaa      1610
113 atctgttaga aaagactttg aaatacttca gtgcaaagtg tgtgtgtgtg aagtttagtt      1670
114 atatcttcat cttcagatga agttttaaag cactttgtag ttctctattg ccaacaattt      1730
115 aatgtttatg tgttgccaat tcttgcaacc actgccctac caaacctgtg ggttgcaaat      1790
116 cagaactaaa attctaagca cgtttcaaag atgaacactt ttgttaagac ccctattgcc      1850
117 tcttcttcat gctcattttt tacttttttt aaaaggtact tttctcatca cattgtagag      1910
118 aggtctgcat tctcattgga aatgtctgtt tagctttata aaacaaacac tttgctgaaa      1970
119 taggaaaatg agccttattg acaattaagt gcttcttgca gcagggtggtc aaagaaaagc      2030
120 atgactaata cgacctatta gagtaatcta catctggacc attccttaag ttttccctca      2090
121 ccgacagtac catcatgcct tgagtgttct tttctcccaa gtgctattcc ttaaacacga      2150
122 gagttttacca gttgccta atgcaataa aaaatgcttt gagatagcta actgcccata      2210

```

RAW SEQUENCE LISTING

DATE: 05/07/2002

PATENT APPLICATION: US/09/874,162A

TIME: 14:55:00

Input Set : A:\05311-024001.txt

Output Set: N:\CRF3\05072002\I874162A.raw

```

123 aaacaaactc aaattgctta taaagtttct tcccatgttc ccatttgatg aaaagtctta 2270
124 catcacatat aactgggaag caggggtccc tcctcaattt tcagacattt tgaaaggatg 2330
125 acagttctgt ttgttagatg agtaaacctc tatattcata agttctaaaa tccttcatta 2390
126 tgagggattc aaagtattta taaaaacact gccctctaaa aatttcctca gatctgaagt 2450
127 atggtcttgg tcctgaatat acagtgttat cctatgttta aaagggatg ccagacatga 2510
128 gacgcaacta gttggtgcat aagaaggccc cacttggcta tttcatatct acctacaatt 2570
129 gaccaaaaaa aatttttttag gccagcaatt attatttagc ttcgctcttt ctagtgcaag 2630
130 aaactgcagg ctggatcagt agttcaacag ctaaacagtc ataaaatagt cattgtgcat 2690
131 gttaaatttc tttcaatgct ttcaaagata aattccaatt tctatttact tattcattgt 2750
132 gacagtatta ctaaacagggt aaggatggga atattttgtt atactgtgta tagtgaatgt 2810
133 attgtactgt gtctgtgaaa actgtgcttt aaattatatt ttcatatgtt ttgttgggga 2870
134 cagagcacat taagtctgaa agcaacagag gtttgtttta gaactgaagg caatttaatc 2930
135 aaaattcctg tcaagaaaag ctgcttataa atgtaaataa aatcacattt aaaataaact 2990
136 gcctctgacc caaaaataaa 3010

```

138 <210> SEQ ID NO: 2

139 <211> LENGTH: 243

140 <212> TYPE: PRT

141 <213> ORGANISM: Homo sapiens

143 <400> SEQUENCE: 2

```

144 Met Thr Gly Ile Ala Ala Ala Ser Phe Phe Ser Asn Thr Cys Arg Phe
145 1 5 10 15
146 Gly Gly Cys Gly Leu His Phe Pro Thr Leu Ala Asp Leu Ile Glu His
147 20 25 30
148 Ile Glu Asp Asn His Ile Asp Thr Asp Pro Arg Val Leu Glu Lys Gln
149 35 40 45
150 Glu Leu Gln Gln Pro Thr Tyr Val Ala Leu Ser Tyr Ile Asn Arg Phe
151 50 55 60
152 Met Thr Asp Ala Ala Arg Arg Glu Gln Glu Ser Leu Lys Lys Lys Ile
153 65 70 75 80
154 Gln Pro Lys Leu Ser Leu Thr Leu Ser Ser Val Ser Arg Gly Asn
155 85 90 95
156 Val Ser Thr Pro Pro Arg His Ser Ser Gly Ser Leu Thr Pro Pro Val
157 100 105 110
158 Thr Pro Pro Ile Thr Pro Ser Ser Phe Arg Ser Ser Thr Pro Thr
159 115 120 125
160 Gly Ser Glu Tyr Asp Glu Glu Glu Val Asp Tyr Glu Glu Ser Asp Ser
161 130 135 140
162 Asp Glu Ser Trp Thr Thr Glu Ser Ala Ile Ser Ser Glu Ala Ile Leu
163 145 150 155 160
164 Ser Ser Met Cys Met Asn Gly Gly Glu Glu Lys Pro Phe Ala Cys Pro
165 165 170 175
166 Val Pro Gly Cys Lys Lys Arg Tyr Lys Asn Val Asn Gly Ile Lys Tyr
167 180 185 190
168 His Ala Lys Asn Gly His Arg Thr Gln Ile Arg Val Arg Lys Pro Phe
169 195 200 205
170 Lys Cys Arg Cys Gly Lys Ser Tyr Lys Thr Ala Gln Gly Leu Arg His
171 210 215 220
172 His Thr Ile Asn Phe His Pro Pro Val Ser Ala Glu Ile Ile Arg Lys
173 225 230 235 240

```

RAW SEQUENCE LISTING

DATE: 05/07/2002

PATENT APPLICATION: US/09/874,162A

TIME: 14:55:00

Input Set : A:\05311-024001.txt

Output Set: N:\CRF3\05072002\I874162A.raw

```

174 Met Gln Gln
177 <210> SEQ ID NO: 3
178 <211> LENGTH: 729
179 <212> TYPE: DNA
180 <213> ORGANISM: Homo sapiens
182 <400> SEQUENCE: 3
183 atgacaggca tgcgcgcgc ctccttcttc tccaatacct gccgattcgg gggctgcgga      60
184 ctccacttcc ccaccctggc cgacctcatc gagcacatcg aggacaacca catcgataca      120
185 gatccacggg ttttagaaaa acaagaatta cagcagccaa cctatgttgc cctgagttac      180
186 ataaatagat tcatgacaga tgctgcccgc cgagagcagg agtccctaaa gaagaagatt      240
187 cagccgaagc tctcgttgac tctgtccagc tcagtgtctc gagggaatgt gtccactccc      300
188 ccacgccaca gcagtggaag ccttactccc cccgtgacct caccatcac cccctcctct      360
189 tcattccgca gcagcactcc gacaggcagc gagtatgacg aggaggagggt ggactatgag      420
190 gagtccgaca gcgatgagtc ctggaccaca gagagtgccg tcagctccga agccatcctc      480
191 agctccatgt gcatgaatgg aggggaagag aagccttttg cctgccagt tcctggatgt      540
192 aaaaagagat acaagaatgt gaatggcata aagtatcacg ctaagaatgg tcacagaaca      600
193 cagattcgtg tccgcaaacc attcaagtgt cgctgtggga agagttacaa gacagctcag      660
194 ggcctgcggc accacacaat caatttccat ccccggtgtg cggctgagat tatcaggaag      720
195 atgcagcaa                                         729
197 <210> SEQ ID NO: 4
198 <211> LENGTH: 4441
199 <212> TYPE: DNA
200 <213> ORGANISM: Homo sapiens
202 <220> FEATURE:
203 <221> NAME/KEY: CDS
204 <222> LOCATION: (195)...(2411)
206 <400> SEQUENCE: 4
207 ctctgaggag acactttttt tttcctccct ccttccctcc tctcctcctc ccttcccttc      60
208 cctctcctc cctctctctc tcttccccc ctcggtccgc cggagcctgc tggggcgagc      120
209 ggttggtatt gcaggcgctt gctctccggg gccgcccggc gggtagctgg cggggggagg      180
210 aggcaggaac cgcg atg gcg cct cag aag cac gcc ggt ggg gga ggg gcc      230
211 Met Ala Pro Gln Lys His Gly Gly Gly Gly Gly Gly
212 1 5 10
214 ggc tgc ggg ccc agc gcg ggg tcc ggg gga gcc ggc ttc ggg ggt tgc      278
215 Gly Ser Gly Pro Ser Ala Gly Ser Gly Gly Gly Gly Phe Gly Gly Ser
216 15 20 25
218 gcg gcg gtg gcg gcg gcg acg gct tgc gcc gcc aaa tcc gcc gcc ggg      326
219 Ala Ala Val Ala Ala Ala Thr Ala Ser Gly Gly Lys Ser Gly Gly Gly
220 30 35 40
222 agc tgt gga ggg ggt gcc agt tac tgc gcc tcc tcc tcc tcc tcc gcc      374
223 Ser Cys Gly Gly Gly Gly Ser Tyr Ser Ala Ser Ser Ser Ser Ser Ala
224 45 50 55 60
226 gcg gca gcg gcg ggg gct gcg gtg tta ccg gtg aag aag ccg aaa atg      422
227 Ala Ala Ala Ala Gly Ala Ala Val Leu Pro Val Lys Lys Pro Lys Met
228 65 70 75
230 gag cac gtc cag gct gac cac gag ctt ttc ctc cag gcc ttt gag aag      470
231 Glu His Val Gln Ala Asp His Glu Leu Phe Leu Gln Ala Phe Glu Lys
232 80 85 90
234 cca aca cag atc tat aga ttt ctt cga act cgg aat ctc ata gca cca      518

```

RAW SEQUENCE LISTING

DATE: 05/07/2002

PATENT APPLICATION: US/09/874,162A

TIME: 14:55:00

Input Set : A:\05311-024001.txt

Output Set: N:\CRF3\05072002\I874162A.raw

235	Pro	Thr	Gln	Ile	Tyr	Arg	Phe	Leu	Arg	Thr	Arg	Asn	Leu	Ile	Ala	Pro	
236			95					100					105				
238	ata	ttt	ttg	cac	aga	act	ctt	act	tac	atg	tct	cat	cga	aac	tcc	aga	566
239	Ile	Phe	Leu	His	Arg	Thr	Leu	Thr	Tyr	Met	Ser	His	Arg	Asn	Ser	Arg	
240		110					115					120					
242	aca	aac	atc	aaa	agg	aaa	aca	ttt	aaa	gtt	gat	gat	atg	tta	tca	aaa	614
243	Thr	Asn	Ile	Lys	Arg	Lys	Thr	Phe	Lys	Val	Asp	Asp	Met	Leu	Ser	Lys	
244	125					130					135					140	
246	gta	gag	aaa	atg	aaa	gga	gag	caa	gaa	tct	cat	agc	ttg	tca	gct	cat	662
247	Val	Glu	Lys	Met	Lys	Gly	Glu	Gln	Glu	Ser	His	Ser	Leu	Ser	Ala	His	
248				145						150					155		
250	ttg	cag	ctt	acg	ttt	act	ggt	ttc	ttc	cac	aaa	aat	gat	aag	cca	tca	710
251	Leu	Gln	Leu	Thr	Phe	Thr	Gly	Phe	Phe	His	Lys	Asn	Asp	Lys	Pro	Ser	
252			160						165					170			
254	cca	aac	tca	gaa	aat	gaa	caa	aat	tct	gtt	acc	ctg	gaa	gtc	ctg	ctt	758
255	Pro	Asn	Ser	Glu	Asn	Glu	Gln	Asn	Ser	Val	Thr	Leu	Glu	Val	Leu	Leu	
256			175				180					185					
258	gtg	aaa	gtt	tgc	cac	aaa	aaa	aga	aag	gat	gta	agt	tgt	cca	ata	agg	806
259	Val	Lys	Val	Cys	His	Lys	Lys	Arg	Lys	Asp	Val	Ser	Cys	Pro	Ile	Arg	
260		190				195					200						
262	caa	gtt	ccc	aca	ggt	aaa	aag	cag	gtg	cct	ttg	att	cct	gac	ctc	aat	854
263	Gln	Val	Pro	Thr	Gly	Lys	Lys	Gln	Val	Pro	Leu	Ile	Pro	Asp	Leu	Asn	
264	205				210				215						220		
266	caa	aca	aaa	ccc	gga	aat	ttc	ccg	tcc	ctt	gca	gtt	tcc	agt	aat	gaa	902
267	Gln	Thr	Lys	Pro	Gly	Asn	Phe	Pro	Ser	Leu	Ala	Val	Ser	Ser	Asn	Glu	
268				225					230				235				
270	ttt	gaa	cct	agt	aac	agc	cat	atg	gtg	aag	tct	tac	tcg	ttg	cta	ttt	950
271	Phe	Glu	Pro	Ser	Asn	Ser	His	Met	Val	Lys	Ser	Tyr	Ser	Leu	Leu	Phe	
272			240					245				250					
274	aga	gtg	act	cgt	cca	gga	aga	aga	gag	ttt	aat	gga	atg	att	aat	gga	998
275	Arg	Val	Thr	Arg	Pro	Gly	Arg	Arg	Glu	Phe	Asn	Gly	Met	Ile	Asn	Gly	
276			255				260				265						
278	gaa	acc	aat	gaa	aat	att	gat	gtc	aat	gaa	gag	ctt	cca	gcc	aga	aga	1046
279	Glu	Thr	Asn	Glu	Asn	Ile	Asp	Val	Asn	Glu	Glu	Leu	Pro	Ala	Arg	Arg	
280		270				275					280						
282	aaa	cga	aat	cgt	gag	gat	ggg	gaa	aag	aca	ttt	gtt	gca	caa	atg	aca	1094
283	Lys	Arg	Asn	Arg	Glu	Asp	Gly	Glu	Lys	Thr	Phe	Val	Ala	Gln	Met	Thr	
284	285			290					295						300		
286	gta	ttt	gat	aaa	aac	agg	cgc	tta	cag	ctt	tta	gat	ggg	gaa	tat	gaa	1142
287	Val	Phe	Asp	Lys	Asn	Arg	Arg	Leu	Gln	Leu	Leu	Asp	Gly	Glu	Tyr	Glu	
288				305				310					315				
290	gta	gcc	atg	cag	gaa	atg	gaa	gaa	tgt	cca	ata	agc	aag	aaa	aga	gca	1190
291	Val	Ala	Met	Gln	Glu	Met	Glu	Glu	Cys	Pro	Ile	Ser	Lys	Lys	Arg	Ala	
292			320					325					330				
294	aca	tgg	gag	act	att	ctt	gat	ggg	aag	agg	ctg	cct	cca	ttc	gaa	aca	1238
295	Thr	Trp	Glu	Thr	Ile	Leu	Asp	Gly	Lys	Arg	Leu	Pro	Pro	Phe	Glu	Thr	
296			335				340				345						
298	ttt	tct	cag	gga	cct	acg	ttg	cag	ttc	act	ctt	cgt	tgg	aca	gga	gag	1286
299	Phe	Ser	Gln	Gly	Pro	Thr	Leu	Gln	Phe	Thr	Leu	Arg	Trp	Thr	Gly	Glu	

VERIFICATION SUMMARY

DATE: 05/07/2002

PATENT APPLICATION: US/09/874,162A

TIME: 14:55:01

Input Set : A:\05311-024001.txt

Output Set: N:\CRF3\05072002\I874162A.raw